Math 203 Reading Notes

David Stein, Spring 2014

Section 1.1: Systems of Linear Equations

A linear equation has the form: $a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$, where a_n, b are real or complex numbers and x_n are variables

A system of linear equations is a collection of one or more linear equations:
$2x_1 - x_2 + 1.5x_3 = 8$
$x_1 - 4x_3 = -7$

A solution to a system is a list of numbers $s_1, s_2, \ldots s_n$ that, when substituted in for $x_1, x_2, \ldots x_n$, produce a set of valid statements for a particular system

The solution set of a linear system is the set of all possible solutions - linear systems are equivalent if they have the same solution sets

A particular system of linear equations can be represented as a matrix, with each row representing one of the equations - can be a coefficient matrix (with only the coefficients $a_1, \ldots$), or an augmented matrix (also including b)

Example:
$x_1 - 2x_2 + x_3 = 0$
$2x_2 - 8x_3 = 8$
$-4x_1 + 5x_2 + 9x_3 = -9$

Coefficient matrix: $\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix}$

Augmented matrix: $\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix}$

Either type of matrix can be manipulated in three ways to produce an equivalent matrix that represents the same system:

- Interchanging the position of any two rows:
$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 1 \\ -4 & 5 & 9 \\ 0 & 2 & -8 \end{bmatrix}$

- Multiplying a row by a constant number:
$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix} \sim \begin{bmatrix} 2 & -4 & 2 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix}$

- Subtracting a multiple of one row from another row:
$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 1 \\ 1 & 0 & -7 \\ -4 & 5 & 9 \end{bmatrix}$

These forms of manipulation can be used to produce a simpler, but still equivalent, system of equations:

System $A : x_1 - 2x_2 + x_3 = 0; 2x_2 - 8x_3 = 8; -4x_1 + 5x_2 + 9x_3 = -9$

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ -4 & 5 & 9 & -9 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & -3 & 13 & -9 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix} \sim$$
$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 & 32 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

System $B : x_1 = 29; x_2 = 16; x_3 = 3$

These two systems are equivalent ($A \sim B$), but B is simpler than A - also, all of these matrices are row-equivalent with one another - any two row-equivalent matrices have the same solution set

Some systems may have no solutions: $x_2 - 4x_3 = 8$
$2x_1 - 3x_2 + 2x_3 = 1$
$5x_1 - 8x_2 + 7x_3 = 1$
Augmented matrix: $\begin{bmatrix} 0 & 1 & -4 & 8 \\ 2 & -3 & 2 & 1 \\ 5 & -8 & 7 & 1 \end{bmatrix} \sim \ldots \sim \begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & \frac{5}{2} \end{bmatrix}$

The last row creates a problem: $0x_1 + 0x_2 + 0x_3 = \frac{5}{2}$ - for any system with an augmented matrix presenting a row with zero coefficients and a nonzero sum, the system has no solutions and is "inconsistent" - any system with at least one solution is "consistent"

Two questions frequently arise with these types of matrices:

1. Does at least one solution exist? (i.e., is the system consistent?)
2. If so, does *only* one solution exist?

Section 1.2: Row Reduction and Echelon Form

The three types of transformation rules can be applied to transform a matrix into echelon form, with the following properties:

1. All nonzero rows are above rows of all zeros
2. The first nonzero entry in each row (the "pivot" of the row) is in a column to the right of the pivot of the row above it
3. The entries in all rows below a pivot are zero

Example of echelon form: $\begin{bmatrix} 4 & 2 & 0 & 1 & 3 \\ 0 & 3 & 2 & 2 & 5 \\ 0 & 0 & 0 & 2 & 1 \end{bmatrix}$

Further transformation can be applied to produce a matrix in reduced echelon form with the following additional properties:

4. The pivot in every row is 1
5. Each pivot is the only nonzero number in its column

Example of reduced echelon form: $\begin{bmatrix} 1 & 0 & 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & 1 & 4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

Every matrix can be reduced to exactly one reduced echelon matrix

Row reduction algorithm: Any matrix can be transformed to echelon form as follows:

1. Interchange all zero rows to the bottom
2. Find the pivot in the top row (optionally, scale the row into a pivot value of 1, and/or interchange another row into its place)
3. Subtract the top row from every row below it to make the pivot column 0 below the pivot
4. Do the same for each successive row down
5. Repeat steps 2-4 for succeeding rows

Further processing can transform the matrix into reduced echelon form:

6. Starting at the bottom, for every row with a pivot, subtract the row from each row above it to make the pivot the only nonzero entry in the column
7. Scale every row to make its pivot 1

A reduced echelon form is the minimal equivalent form of a system of linear equations, in which every nonzero row specifies the value of a single variable:
$A = \begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 2 & 6 \end{bmatrix}$ is equivalent to: $x_1 = 4; x_2 + 2x_3 = 6$

Every column having a pivot defines a "basic" variable - every column with no pivot, such as x_3, defines a "free" variable that can have any value in this system of equations - the "general solution" of a system is the description of the value of each variable (i.e., each variable is specified in "parametric" form):

$$A = \begin{cases} x_1 = 4 \\ x_2 = 6 - 2x_3 \\ x_3 \text{ is free} \end{cases}$$

A solution set with only basic variables has exactly one solution - a solution set with at least one free variable has infinitely many solutions (i.e.,

Section 1.3: Vector Equations

An $m \times n$ matrix can be geometrically conceptualized as a set of n columns that each represents a vector in R^m space - a single-column matrix represents a single vector - zero vector: any vector with all entries set to zero

Vector addition: $\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1+3 \\ 2+4 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$

Scalar multiplication: $u = \begin{bmatrix} 2 & 3 \end{bmatrix}; c = 4; cu = \begin{bmatrix} 2 \cdot 4 \\ 3 \cdot 4 \end{bmatrix} = \begin{bmatrix} 8 \\ 12 \end{bmatrix}$

Properties of vector math:

- u + 0 = 0 + u = u
- u + (-u) = 0
- Commutative: $u + v = v + u$
- Associative: $(u + v) + w = u + (v + w)$
- Distributive sum: $c(u + v) = cu + cv$
- Distributive multiplication: $c(du) = d(cu) = (cd)u$

A linear system can represent an equation of vectors:
$c_1v_1 + c_2v_2 + \ldots + c_pv_p = y$ (where y is called the linear combination of these vectors) - the solution for this system is the entire set of valid weights that causes the vectors to sum to y, i.e., the entire set of valid linear combinations of these vectors - example:
$a_1 = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}; a_2 = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}; b = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$
$x_1a_1 + x_2a_2 = b \rightarrow$ Given vectors a_1, a_2, what combination of weights for each vector produces a vector sum equal to b? - that is: $x_1 \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$

This can be solved with an augmented matrix in reduced echelon form:

$$\begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 7 \\ 0 & 9 & 18 \\ 0 & 16 & 32 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 7 \\ 0 & 1 & 2 \\ 0 & 16 & 32 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Solution: $3 \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + 2 \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$

If, in the solution set, none of the weights is a free variable, then there is at most one solution; if at least one weight is free, then there are infinitely many solutions - the number of free variables defines whether the solution set defines a point, a line, a plane, etc.

For a particular set of vectors $a_1, a_2, \ldots a_n$ and weights $x_1, x_2, \ldots x_n$, the "span" of the vector system is the entire set of linear combinations of $x_1 a_1 + x_2 a_2 + \ldots x_n a_n$

Geometrically conceptualized: For vectors in R^3, each vector equation defines a plane (may be a line if it only includes one variable) - the span is the intersection of all of these planes - for a consistent vector system with no free variables, the span is a point; one free variable = a line; two free variables = a plane; etc. - an inconsistent vector system has no intersection

Section 1.4: The Matrix Equation Ax = b

Matrix equation: The product of a matrix and a vector produces a new matrix:

$$\begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 a_1 + x_2 a_2 + x_3 a_3$$

Can be extended to multiple dimensions: $\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_1 x_1 + a_2 x_1 \\ a_1 x_2 + a_2 x_2 \end{bmatrix}$ - note that the resulting matrix is simply the augmented matrix

An equation $Ax = b$ has a solution if and only if b is a linear combination of the columns of A (i.e., if b is in Span(A)) - another question: what is the set of b

in the span of A? - if every b has a solution of A, then we say that A spans R^m

Some equivalent statements about an $A = m \times n$ matrix (either all true or all false):

1. The columns of A span R^m.
2. A has a pivot position in every row.
3. For each b in R^m, the equation $Ax = b$ has a solution.
4. Each b in R^m is a linear combination of the columns of A.

Properties of matrix equations, where A is a matrix, u, v are vectors, and c is a scalar constant:

- $A(u + v) = Au + Av$
- $A(cu) = c(Au)$

Section 1.5: Solution Sets of Linear Systems

Homogeneous system: A linear system that can be rewritten as $Ax = 0$ - these systems always have one solution, where x is the zero vector (this is the "trivial" solution), but may also have others where x is a nonzero vector ("nontrivial" solutions) - this only occurs where A has at least one free variable

Example:
$3x_1 + 5x_2 - 4x_3 = 0$
$-3x_1 - 2x_2 + 4x_3 = 0$
$6x_1 + x_2 - 8x_3 = 0$
Solution: $\begin{bmatrix} 3 & 5 & -4 & 0 \\ -3 & -2 & 4 & 0 \\ 6 & 1 & -8 & 0 \end{bmatrix} \sim \begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & -9 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$
This matrix has a free variable, and thus has nontrivial solutions.

A single linear equation can be treated as a system:
$10\text{x}_1 - 3x_2 - 2x_3 = 0$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.3x_2 + 0.2x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.3x_2 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.2x_3 + 0 + x_3 \end{bmatrix} = x_2 \begin{bmatrix} 0.3 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0.2 \\ 0 \\ 1 \end{bmatrix} \text{ (where } x_2, x_3 \text{ are free)}$$

Parametric vector form: A solution of the form $x = p + x_1 u + x_2 v + \ldots$, where p, u, v are vectors - p is needed only for nonhomogeneous systems

Example: Describe all solutions for the system $Ax = b$, where:

$$A = \begin{bmatrix} 3 & 5 & -4 \\ -3 & -2 & 4 \\ 6 & 1 & -8 \end{bmatrix}, B = \begin{bmatrix} 7 \\ -1 \\ -4 \end{bmatrix}$$

Answer: $\begin{bmatrix} 3 & 5 & -4 & 7 \\ -3 & -2 & 4 & -1 \\ 6 & 1 & -8 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -\frac{4}{3} & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

General Solution: $\begin{cases} x_1 = -1 + \frac{4}{3}x_3 \\ x_2 = 2 \\ x_3 \text{ is free} \end{cases}$

Parametric Vector Form: $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix}$

Section 1.6: Applications of Linear Systems

Linear systems can be used to model an economy - as an example, consider this table, which shows how many resources are used by each industry to produce one unit of a good:

Purchased By →	Coal	Electric	Steel
Coal	0.0	0.4	0.6
Electric	0.6	0.1	0.2
Steel	0.4	0.5	0.2

This model (an "exchange model") can be used to determine how much of each good should be manufactured in an equilibrium system:

$$\begin{bmatrix} 1.0 & -0.4 & -0.6 & 0.0 \\ -0.6 & 0.9 & -0.2 & 0.0 \\ -0.4 & -0.5 & 0.8 & 0.0 \end{bmatrix} \sim \begin{bmatrix} 1.0 & 0.0 & -0.94 & 0.0 \\ 0 & 1.0 & -0.85 & 0.0 \\ 0 & 0 & 0 & 0.0 \end{bmatrix}$$

$$\text{Solution: } \begin{cases} q_C = 0.94 q_E \\ q_E = 0.85 q_S \\ q_S \text{ is free} \end{cases}$$

Chemistry: Linear algebra can be used to determine reaction products:

$(x_1)\ \mathrm{C}_3 H_8 + (x_2)\ \mathrm{O}_2 \to (x_3)\ \mathrm{CO}_2 + (x_4)\ \mathrm{H}_2 O$

Construct a matrix with the rows representing C, H, and O:

$$x_1 \begin{bmatrix} 3 \\ 8 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$$

$$x_1 \begin{bmatrix} 3 \\ 8 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} - x_3 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} - x_4 \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\text{Solution: } \begin{cases} x_1 = \frac{1}{4} x_4 \\ x_2 = \frac{5}{4} x_4 \\ x_3 = \frac{3}{4} x_4 \\ x_4 \text{ is free} \end{cases}$$

Network flow (e.g., traffic, or water in pipes) - consider the following map:

- Intersection A: 800 units flowing in; flowing out to B (rate x_2) and D (rate x_1)
- Intersection B: Flowing in from A and C; flowing out at the rate of 300 + x_3
- Intersection c: 500 units flowing in; flowing out to B (rate x_4 and D (rate x_5)

- Intersection D: 600 units flowing out

This network flow model can be solved with a simple linear equation:

Intersection	Inflow	Outflow
A	800	$x_1 + x_2$
B	$x_2 + x_4$	300 + x_3
C	500	$x_4 + x_5$
D	$x_1 + x_5$	600

$$\sim \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 800 \\ 0 & 1 & -1 & 1 & 0 & 300 \\ 0 & 0 & 0 & 1 & 1 & 500 \\ 1 & 0 & 0 & 0 & 1 & 600 \end{bmatrix}$$

$$\text{Solution: } \begin{cases} x_1 = 600 - x_5 \\ x_2 = 200 + x_5 \\ x_3 = 400 \\ x_4 = 500 - x_5 \\ x_5 \text{ is free} \end{cases}$$

Section 1.7: Linear Independence

Section 1.5 studied homogeneous systems from the perspective of the coefficients - however, homogeneous systems also have important details involving the relationships of the vectors: does a homogeneous vector system have only the trivial solution (are the vectors linearly independent) or more than one solution (are the vectors linearly dependent)? - this question can be represented as a vector equation:

$$x_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Question: Does a nonzero set (x_1, x_2, x_3) exist?

...or as a matrix equation:

$$\begin{bmatrix} 1 & 4 & 2 & 0 \\ 2 & 5 & 1 & 0 \\ 3 & 6 & 0 & 0 \end{bmatrix}$$

Determining linear independence is simple, and can often be done by inspection:

- Any system that contains a zero vector is always linearly dependent
- Any system with more vectors than entries is always linearly dependent: at least one variable must be free
- A system with only one vector is linearly dependent if the vector is the zero vector (since any coefficient works), and linearly independent if the vector is nonzero (since the only valid coefficient is zero)
- A system of two vectors is linearly dependent if one vector is a multiple of the other, and linearly independent if not
- A system with two or more vectors is linearly dependent if one vector is a linear combination of the others, and linearly independent if not

Section 1.8: Introduction to Linear Transformations

Matrix multiplication (i.e., a matrix times a vector) can be viewed as a technique for transforming an input vector to an output vector according to the matrix - i.e., a function T maps any vector in an input space in R^n to an output space R^m, according to an $m \times x$ matrix - dimensions need not be the same: e.g., three-dimensional projection onto two-dimensional surface

Some terminology:

- T is called a function, a transformation, or a mapping
- R^n is called the domain of T
- R^m is called the codomain of T
- For each x in $R^n, T(x)$ is called the image of x in R^m
- The set of all images x in R^m is called the range of T

Properties of linear transformations (for a linear function T, vectors u, v, and scalar constants c, d):

- $T(0) = 0T(u+v) = T(u) + T(v)$
- $T(cu) = cT(u)$
- $T(cu+dv) = T(cu) + T(dv)$

Section 1.9: The Matrix of a Linear Transformation

For every linear transformation $R^n \to R^n$, there exists a unique matrix A such that $T(x) = Ax$ for all x in R^n - A is called the "standard matrix" for T

Given a particular function T, we can figure out the entries in each row by feeding in each column of an identity matrix and examining its image - example for a given function T:

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \to T(e_1) = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$$

$$e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \to T(e_2) = \begin{bmatrix} 4 & 5 & 6 \end{bmatrix}$$

$$e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \to T(e_3) = \begin{bmatrix} 7 & 8 & 9 \end{bmatrix}$$

$$\therefore A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

Some common transformations:

- Scale an input vector by a constant c : $A = \begin{bmatrix} c & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & c \end{bmatrix}$ - matrices with

c less than 1 are called a contraction; matrices with c greater than 1 are called a dilation

- Reflection over x-axis: $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
- Reflection over the line $y = x : A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
- Reflection over the origin: $A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$
- Rotating a point around the origin counterclockwise by an angle $\phi : A = \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix}$
- Horizontal shear: $A = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$
- Projection onto x-axis: $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

Some properties of mappings:

- A mapping T of R^n is said to map **onto** R^m if every b in R^m is the image of at least one x of R^n (i.e., every location in the codomain is mappable via T)
- A mapping T of $R^n \to R^m$ is called $one - to - one$ if every b in R^m is the image of **at most** one x in R^n - this is true if and only if the columns of A are linearly independent

Section 2.1: Matrix Operations

The main diagonal of a matrix is the set of entries from top-left to bottom-right

Diagonal matrix: Square matrix whose non-diagonal entries are zero:
$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Zero matrix: Any matrix where all entries are 0

Matrix sum: Two same-size matrices can be added on a per-entry basis:
$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$

Equal matrices: Two or more matrices with same size and identical entries

Scalar multiple: Multiplying any matrix by a scalar = multiplying the entries by the scalar:
$2 \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}$

Matrix multiplication: Two matrices can be multiplied: multiply the entire first matrix by each column of the second matrix - i.e., the resulting matrix is a linear combination of the first matrix with each column of the second matrix
$\begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} \mid \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 13 \\ 17 \\ 21 \end{bmatrix} \mid \begin{bmatrix} 18 \\ 24 \\ 30 \end{bmatrix} = \begin{bmatrix} 13 & 18 \\ 17 & 24 \\ 21 & 30 \end{bmatrix}$

Properties of matrix multiplication:

- If multiplying $A \times B$, the resulting matrix is written AB
- Size requirements: The width of the first matrix equals the height of the second matrix: i.e., left matrix is $a \times b$; right matrix is $b \times c$ - the resulting matrix is $a \times c$ (see above)
- Matrix multiplication is associative: $A(BC) = (AB)C$
- Matrix multiplication is **not** commutative: $AB \neq BA$
- Matrix multiplication is distributive: $A(B + C) = AB + AC$, and $(A + B)C = AC + BC$
- Scalar distribution: $r(AB) = (rA)B = A(rB)$
- Multiplying a matrix by a (square) identity matrix returns the original matrix: $I_m A = AI_n = A$
- Cancellation laws do not apply: $AB = AC$ does not imply that $B = C$

- If $AB =$ zero matrix, it's not necessarily true that at least one of A, B is the zero matrix: $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

Transpose of a matrix: Simply mirror over minor diagonal:
$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}; A^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$

Properties of transpose:

- $(A^T)^T = A$
- $(A + B)^T = A^T + B^T$
- $(rA)^T = rA^T$
- $(AB)^T = B^T A^T$

Section 2.2: The Inverse of a Matrix

Some square matrices A have an inverse matrix A^{-1} such that $AA^{-1} = I$, and $A^{-1}A = I$ - example:
$A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}; C = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix} \to AC = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; CA = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
$\therefore C = A^{-1}; A = C^{-1}$.

Properties of matrix inverse:

- Not all matrices have this property: those that do are called non-singular matrices; those that don't are called singular matrices
- For a 2x2 matrix, the inverse can be found using the determinant: If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the determinant $D = ac - bd$ - for a matrix A, if $D = 0$, the matrix is singular; if $D \neq 0$, the matrix is nonsingular, and its inverse $A^{-1} = \frac{1}{D} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ (i.e., switch the entries along the major diagonal, and multiply the entries along the minor diagonal by -1)

- For an invertible matrix A (size $n \times x$), for every b in $R^n \rightarrow Ax = b, x = A^{-1}b$
- For an invertible matrix A, $(A^{-1})^{-1} = A$
- For two invertible $n \times x$ matrices $A, B : (AB)^{-1} = B^{-1}A^{-1}$
- The transpose of an inverse is the same as the inverse of the transpose: $(A^T)^{-1} = (A^{-1})^T$

Elementary matrix: An identity matrix with a single row operation performed on it (e.g., multiplying a row by a number; switching two rows; or adding a multiple of one row to another) - every elementary matrix is invertible

For an $m \times n$ matrix A, performing any single row operation on A can be represented as the matrix multiple EA, where $E =$ an $m \times m$ elementary matrix E with the same row operation

Algorithm for finding the inverse: For any invertible matrix A, horizontally concatenate the identity matrix I, and row-reduce A to reduced echelon form - the same operations performed on the identity matrix produce A^{-1}

Example: $A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$:

$$\begin{bmatrix} 2 & 5 & 1 & 0 \\ -3 & -7 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & \frac{5}{2} & \frac{1}{2} & 0 \\ -3 & -7 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & \frac{5}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{3}{2} & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -7 & -5 \\ 0 & \frac{1}{2} & \frac{3}{2} & 1 \end{bmatrix} \sim$$

$$\begin{bmatrix} 1 & 0 & -7 & -5 \\ 0 & 1 & 3 & 2 \end{bmatrix}$$

$$A^{-1} = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$$

Section 2.3: Characterizations of Invertible Matrices

According to the Invertible Matrix Theorem (IMT), the following statements are all true about invertible (non-singular) matrices, and are all false about singular matrices:

- A is an invertible matrix.
- A^T is an invertible matrix.
- A is row-equivalent to I_n.
- A has n pivot positions.
- The equation $Ax = 0$ has only the trivial solution.
- The columns of A are linearly independent.
- The linear transformation $x \to Ax$ is one-to-one.
- The equation $Ax = b$ has at least one solution for each b in R^n.
- The columns of A span R^n.
- The linear transformation $x \to Ax$ maps R^n onto R^n.
- There is an $n \times n$ matrix $C : CA = I$ and $AC = I$.

Invertible linear transformations: If the matrix A of a linear transformation T is invertible, then there exists another linear transformation S with a matrix A^{-1}, called the inverse of T or T^{-1}, where $S(T(x)) = x$, and $T(S(x)) = x$.

Section 2.4: Partitioned Matrices

A matrix can be partitioned into blocks:
A =

$$\left[\begin{array}{cc|cc} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ \hline 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{array}\right]$$

The individual units can also be written as submatrices:
$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}; A_{11} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}; \ldots$

Matrix addition and scalar multiplication work as expected: just apply to every entry in every submatrix

Matrix multiplication: Multiply each submatrix by each submatrix:
$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}; B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}; AB = \begin{bmatrix} A_{11}B_1 + A_{12}B_2 \\ A_{21}B_1 + A_{22}B_2 \end{bmatrix}$

Example:

$$A = \left[\begin{array}{ccc|cc} 2 & -3 & 1 & 0 & -4 \\ 1 & 5 & -2 & 3 & -1 \\ \hline 0 & -4 & -2 & 7 & -1 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

$$B = \left[\begin{array}{cc} 6 & 4 \\ -2 & 1 \\ -3 & 7 \\ \hline -1 & 3 \\ 5 & 2 \end{array}\right] = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} A_{11}B_1 + A_{12}B_2 \\ A_{21}B_1 + A_{22}B_2 \end{bmatrix}$$

$$A_{11}B_1 = \begin{bmatrix} 15 & 12 \\ 2 & -5 \end{bmatrix}$$

$$A_{12}B_2 = \begin{bmatrix} -20 & -8 \\ -8 & 7 \end{bmatrix}$$

$$A_{11}B_1 + A_{12}B_2 = \begin{bmatrix} -5 & 4 \\ -6 & 2 \end{bmatrix}$$

$$A_{21}B_1 = \begin{bmatrix} 14 & -18 \end{bmatrix}$$

$$A_{22}B_2 = \begin{bmatrix} -12 & 19 \end{bmatrix}$$

$$A_{21}B_1 + A_{22}B_2 = \begin{bmatrix} 2 & 1 \end{bmatrix}$$

$$A_{21}B_2 = AB = \left[\begin{array}{cc} -5 & 4 \\ -6 & 2 \\ \hline 2 & 1 \end{array}\right]$$

Block upper triangle matrix: A partitioned matrix of the form $A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ 0 & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{bmatrix}$

Block diagonal matrix: A partitioned matrix where all blocks off the main diagonal are zero blocks

Section 2.5: Matrix Factorizations

Common problem arising in linear algebra: Given a matrix A and output b, where $Ax = b$, figure out x - this can be difficult if A is large - instead, we can factor A into a product of two simpler matrices:

L, which is a unit lower triangle matrix (e.g.: $\begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}$); and

U, which is an upper triangle (e.g.: $\begin{bmatrix} * & * & * & * & * \\ 0 & * & * & * & * \\ 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$)

If $A = LU$, then $Ax = b$ can be rewritten as $L(Ux) = b$, where each matrix multiplication is easier to evaluate - hence, $Ly = b$, and $Ux = y$; for a given b, we can first solve $Ly = b$ for y, and then solve $Ux = y$ for x

Example: Given $A = \begin{bmatrix} 3 & -7 & -2 & 2 \\ -3 & 5 & 1 & 0 \\ 6 & -4 & 0 & -5 \\ -9 & 5 & -5 & 12 \end{bmatrix}$ and $b = \begin{bmatrix} -9 \\ 5 \\ 7 \\ 11 \end{bmatrix}$, find $x \to Ax = b$.

Solution: First, decompose A into $L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -5 & 1 & 0 \\ -3 & 8 & 3 & 1 \end{bmatrix}$ and $U = \begin{bmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix}$.

Next, figure out $Ly = b$:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ -1 & 1 & 0 & 0 & 5 \\ 2 & -5 & 1 & 0 & 7 \\ -3 & 8 & 3 & 1 & 11 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ 0 & 1 & 0 & 0 & -4 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

$$\to y = \begin{bmatrix} -9 \\ -4 \\ 5 \\ 1 \end{bmatrix}$$

Next, figure out $Ux = y$: $\begin{bmatrix} 3 & -7 & -2 & 2 & -9 \\ 0 & -2 & -1 & 2 & -4 \\ 0 & 0 & -1 & 1 & 5 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & -6 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}$

$\rightarrow x = \begin{bmatrix} 3 \\ 4 \\ -6 \\ -1 \end{bmatrix}$

Some matrices A can be reduced to echelon form without any row interchanges - for these matrices, the following algorithm can be used to factor A into L, U:

1. Prepare L as a lower $n \times n$ triangular matrix (same height as A), and prepare U as a copy of A
2. Choose the first entry of the first column of U - for L, divide every entry in the left column of U by this top value, and copy it into the corresponding spot in L
3. For the first column of U, use row replacement (i.e., subtraction) to zero all of the entries below the pivot
4. Move to the pivot in the next row of U - again, divide the pivot and every entry below it by the pivot, and copy the results into the corresponding entry in L - then, subtract the second row of U from every row below it to zero those entries in U
5. Continue for all following rows of U

Example:

$A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix}$

$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ & 1 & 0 & 0 \\ & & 1 & 0 \\ & & & 1 \end{bmatrix}; U = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix}$

First column: Scale L by 2, then row-substitute U:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & & 1 & 0 \\ -3 & & & 1 \end{bmatrix}; U = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & -9 & -3 & -4 & 10 \\ 0 & 12 & 4 & 12 & -5 \end{bmatrix}$$

Second column: Scale L by 3, then row-substitute U:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ -3 & 4 & & 1 \end{bmatrix}; U = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 4 & 7 \end{bmatrix}$$

Third column (fourth column of U): Scale L by 2, then row-substitute U:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ -3 & 4 & 2 & 1 \end{bmatrix}; U = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

This produces $LU = A$, and any equation $Ax = b$ can be solved more easily by evaluating $L(U(x)) = b$.

For matrices that require row interchanges to reach echelon form, a technique called permuted LU factorization is involved - see study guide

Matrix factorization can be used to calculate the properties of electric circuits - for a circuit having input voltage and current $\begin{bmatrix} v_1 \\ i_1 \end{bmatrix}$ and output voltage and current $\begin{bmatrix} v_2 \\ i_2 \end{bmatrix}$, a transfer matrix A maps the input to the output: $A \begin{bmatrix} v_1 \\ i_1 \end{bmatrix} = \begin{bmatrix} v_2 \\ i_2 \end{bmatrix}$
- for a circuit having a resistor in series, the transfer matrix is $\begin{bmatrix} 1 & -R \\ 0 & 1 \end{bmatrix}$; for a circuit having a resistor in parallel, the transfer matrix is $\begin{bmatrix} 1 & 0 \\ -\frac{1}{R} & 1 \end{bmatrix}$

Section 2.6: The Leontief Input-Output Model

Leontief's work involved an evaluation of a national economy according to linear algebra - an economy with n sectors can be represented as a production vector

x in R^n with the annual output of each sector - in order to produce, each productive sector consumes an amount of each other good, creating “intermediate demand”; and a non-productive portion of the economy (the “open sector”) does not produce goods, but consumes goods according to a “final demand vector” d - the key question is whether there is an a production vector x that is exactly the sum of intermediate demand and final demand

The intermediate each productive sector can be represented as a unit consumption vector, representing the units output by every other sector that are consumed to make one unit from this sector, resulting in a “consumption matrix” C:

	Inputs Consumed Per Unit of Output		
Purchased From	Manufacturing	Agriculture	Services
Manufacturing	0.5	0.4	0.2
Agriculture	0.2	0.3	0.1
Services	0.1	0.1	0.3

(Note: the column sum for each productive sector will typically be less than one; otherwise, the sector destroys value - e.g., if the input and output units are measured in a normalized unit, such as currency, then sectors with a negative column sum are unprofitable)

The consumption matrix reduces the problem to simple linear algebra: $Cx+d = x$, or $(I - C)x = d$

This can be computed over a recursive basis for a number of periods, where the output of each sector every period is used to satisfy the inputs of the other sectors in the period: $x = (I+C+C^2+C^3+\ldots)d$ - finding the result of m rounds of recursion involves calculating C^m - thus, $(I-C)^{-1} \approx I + C + C^2 + \ldots + C^m$ - since all entries are less than 1, C^m rapidly approaches zero, resulting in $Ix^m = d^m$ - this calculation is useful because it reflects how the production x of every sector will change if the final demand d changes: every entry (i, j) in $(I - C)^{-1}$ reflects how much output sector i will have to increase to sustain a unit increase in the final demand of sector j

Section 2.7: Applications to Computer Graphics

Many software tasks involve the manipulation of graphics - e.g., a simple vector graphics object can represent a font for an alphabetic letter as a series of vertices, and the lines connecting them - the font can be manipulated through linear algebra: italics can be represented as a horizontal shear transformation: $\begin{bmatrix} 1 & 0.25 \\ 0 & 1 \end{bmatrix}$; bold can be represented as a horizontal dilation transformation: $\begin{bmatrix} 1.25 & 0 \\ 0 & 1 \end{bmatrix}$

Homogenous coordinates: When rendering R^n graphics using linear algebra operations, a problem is encountered: linear operations always leave the origin fixed (i.e., any manipulation on a point at the origin must leave the point at the origin), but we often want to apply a translation that moves this point - instead, a trick can be used: represent the R^n graphic in R^{n+1}, where the coordinate in the extra dimension is 1 - thus, a two-dimensional graphic with pixels at points (x, y) is represented as a three-dimensional graphic with pixels at points $(x, y, 1)$, and matrix operations can then be applied to achieve translation and other transformations

Useful 2D transformations:

- Translating coordinate (x, y) by $(h, v) : T = \begin{bmatrix} 1 & 0 & h \\ 0 & 1 & v \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x+h \\ y+k \\ 1 \end{bmatrix}$
- Rotating coordinate (x, y) about the origin by $\phi : R_1 = \begin{bmatrix} \cos(\phi) & -\sin(\phi) & 0 \\ \sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{bmatrix}$
- Reflecting coordinate (x, y) over the line $y = x : R_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
- Scaling coordinate (x, y) by $(s, t) : S = \begin{bmatrix} s & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Multiple transformations can be applied in series: $T(R_1(R_2(S(x,y))))$ - such cases just involve repeated matrix multiplication

Homogeneous coordinates can be extended to R^3 coordinates, which are represented as (x,y,z,h) - rendering R^3 coordinates onto a two-dimensional display requires selecting a viewing plane, and then projecting the points of the object onto the viewing plane using a perspective projection mapping (x,y,z) to $(x^*,y^*,0)$

Section 3.1: Determinants

Recalling that for a 2x2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the determinant $D = ad - bc$, and the matrix is invertible if and only if $D \neq 0$

The concept of a determinant can be extended from 2x2 matrices to larger square matrices - for a 3x3 matrix: $\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$, the determinant $D = aei + bfg + cdh - afh - bdi - ceg$ - again, a 3x3 matrix is invertible if and only if $D \neq 0$

A recursive definition for determinants involves breaking down a large matrix into the top row and various submatrices beneath the top row - first, for a square matrix $A = [a_{ij}]$ of size $n \geq 2$, define A_{ij} as the determinant of the matrix formed by deleting row i and column j - example:

$$A = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{bmatrix}$$

$$A_{11} = \begin{bmatrix} f & g & h \\ j & k & l \\ n & o & p \end{bmatrix}$$

$$A_{22} = \begin{bmatrix} a & c & d \\ i & k & l \\ m & o & p \end{bmatrix}$$

Given this definition - for a square matrix $A = [a_{ij}]$ of size $n \geq 2$, the determinant is the sum of n terms of the form $\pm a_{ij}\text{Det } A_{ij}$, with signs alternating - i.e.: det $A = a_{11}\text{Det } A_{11} - a_{12}\text{Det } A_{12} + \ldots + (-1)^{n+1}a_{1n}\text{Det } A_{1n}$

Example:

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

Det $A = a_{11}\text{Det } A_{11} - a_{12}\text{Det } A_{12} + a_{13}\text{Det } A_{13}$

$$= 1 \cdot \text{Det } \begin{bmatrix} 4 & -1 \\ -2 & 0 \end{bmatrix} - 5 \cdot \text{Det } \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix} + 0 \cdot \text{Det } \begin{bmatrix} 2 & 4 \\ 0 & -2 \end{bmatrix}$$

$= 1(0-2) - 5(0-0) + 0(-4-0) = -2$

This can be rewritten as a sum of cofactors, where each cofactor $C_{ij} = (-1)^{i+j}\text{Det } A_{ij}$, such that det $A = a_{11}C_{11} + a_{12}C_{12} + \ldots + a_{1n}C_{1n}$ - this algorithm is called cofactor expansion across the first row of A - can also be done across any row, or down any column

Example:

$$A = \begin{bmatrix} 3 & -7 & 8 & 9 & -6 \\ 0 & 2 & -5 & -7 & 3 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 2 & 4 & -1 \\ 0 & 0 & 0 & -2 & 0 \end{bmatrix}$$

Expanding down the first column of A produces a lot of zeros that simplify our solution:

$$\text{det } A = 3 \cdot \begin{bmatrix} 2 & -5 & 7 & 3 \\ 0 & 1 & 5 & 0 \\ 0 & 2 & 4 & -1 \\ 0 & 0 & -2 & 0 \end{bmatrix} - 0 \cdot C_{21} + 0 \cdot C_{31} - 0 \cdot C_{41} + 0 \cdot C_{51}$$

For the submatrix, again expanding down the first column:

$$\text{det } A = 3 \cdot 2 \cdot \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix} - 0 \cdot C_{21} + 0 \cdot C_{31} - 0 \cdot C_{41}$$

The determinant of this remaining submatrix is easy to calculate as -2, so det $A = -12$.

Repeatedly applying this shortcut - for a triangular matrix, the determinant is just the product of the entries on the main diagonal - also, if any row or column of a matrix is zero, the determinant is necessarily zero

Section 3.2: Properties of Determinants

The effects of elementary row operations on the determinant of a square matrix A:

- Change positions of two rows: det A' = -det A
- Adding one row to another: det A' = det A
- Multiplying one row by a constant k: det $a' = k$ det A

These three properties can be used to transform a matrix in a way that enables a faster calculation of the determinant, especially when combined with the shortcut for a triangular matrix:

$$A = \begin{bmatrix} 1 & -4 & 2 \\ -2 & 8 & 9 \\ -1 & 7 & 0 \end{bmatrix}$$

$$\text{Det} \begin{bmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{bmatrix} = \text{Det} \begin{bmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ -1 & 7 & 0 \end{bmatrix} = \text{Det} \begin{bmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ 0 & 3 & 2 \end{bmatrix} = -\text{Det} \begin{bmatrix} 1 & -4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & -5 \end{bmatrix} =$$

$$-(1 \cdot 3 \cdot (-5)) = 15$$

In general, every invertible matrix can be reduced to echelon form, and then the determinant is just the product of the pivots, factoring in any interchanging and scaling operations performed to get there - as per the Invertible Matrix Theorem, an invertible matrix is linearly independent and has no free variables, and thus has a pivot in every row - if this isn't true, then in at least one row, the entry along the main diagonal is zero, and the determinant is zero; therefore, det $= 0$, and the matrix is not invertible - $\therefore$ a square matrix A is only invertible if det $\neq 0$

By definition, if A is an invertible square matrix, det $A =$ det A^T - this also implies that the effects of elementary row operations on the determinant are identical to elementary *column* operations: i.e., interchanging columns, multiplying columns, or adding one column to another

Multiplication: if A, B are square, invertible matrices of size n, then det $AB =$ (det A) (det B) (NOTE: Addition does not work: det $(A+B) \neq$ det (A) + det (B))

Section 3.3: Cramer's Rule, Volume, and Linear Transformations

Cramer's Rule uses the following notation: For any square matrix A of size n, and for any vector b, $A_i(b)$ is the matrix obtained by replacing column i with b - i.e.: $A_i(b) = [a_1 \dots b \dots a_n]$.

Cramer's Rule: For an invertible matrix A of size n, the unique solution x for any $Ax = b$ is determined by $x_i = \frac{\text{Det } A_i(b)}{\text{Det } A}$ for each $i = 1, \dots n$.

Example: For the system $3x_1 - 2x_2 = 6; -5x_1 + 4x_2 = 8$, find x.
Answer:
$A = \begin{bmatrix} 3 & -2 \\ -5 & 4 \end{bmatrix}; b = \begin{bmatrix} 6 \\ 8 \end{bmatrix}$
Det $A = 2$
$A_1(b) = \begin{bmatrix} 6 & -2 \\ 8 & 4 \end{bmatrix}$; Det $A_1(b) = 40$; $\frac{\text{Det } A_1 b}{\text{Det } A} = 20$
$A_2(b) = \begin{bmatrix} 3 & 6 \\ -5 & 8 \end{bmatrix}$; Det $A_2(b) = 54$; $\frac{\text{Det } A_2 b}{\text{Det } A} = 27$
$x = \begin{bmatrix} 20 \\ 27 \end{bmatrix}$

Laplace transform: This technique represents a system of linear differential equations as a system of linear algebraic equations, with a coefficient involving a parameter

Example:
$3sx_1 - 2x_2 = 4; -6x_1 + sx_2 = 1$

$A = \begin{bmatrix} 3s & -2 \\ -6 & s \end{bmatrix}$

Det $A = 3s^2 - 12 = 3(s+2)(s-2) \rightarrow$ This system has a unique solution only when $s \neq \pm 2$.

$A_1(b) = \begin{bmatrix} 4 & -2 \\ 1 & s \end{bmatrix}$

$x_1 = \frac{\text{Det } A_1(b)}{\text{Det } A} = \frac{4s+2}{3(s+2)(s-2)}$

$A_2(b) = \begin{bmatrix} 3s & 4 \\ -6 & 1 \end{bmatrix}$

$x_2 = \frac{\text{Det } A_2(b)}{\text{Det } A} = \frac{3s+24}{2(s+2)(s-2)} = \frac{s+8}{(s+2)(s-2)}$

Adjugate of A: The transpose of the cofactor matrix of A - example:

$A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix}$

Cofactors of A :

$C_{11} = -2; C_{12} = 3; C_{13} = 5$

$C_{21} = 14; C_{22} = -7; C_{23} = -7$

$C_{31} = 4; C_{32} = 1; C_{33} = -3$

Cofactor matrix of $A = \begin{bmatrix} -2 & 3 & 5 \\ 14 & -7 & -7 \\ 4 & 1 & -3 \end{bmatrix}$

Adjugate of $A =$ transpose of cofactor matrix of $A = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix}$

Cramer's Rule provides an easy way of determining the inverse of a matrix - for any square matrix A of size n, the inverse A^{-1} is the adjugate of A divided by the determinant of A

Two uses of determinants:

- Determinants provide an easy way of determining the area of a parallelogram or parallelpiped (i.e., a three-dimensional parallelogram): given a 2x2 or 3x3 matrix, the area or volume of the shape formed by the vectors is the determinant of the matrix
- Given a 2x2 linear transformation T with standard matrix A and a par-

allelogram S, the area of the parallelogram formed by $T(S)$ is $|\text{Det } A| \cdot$ (area of S). Similarly, given a 3x3 linear transformation T with standard matrix A and a parallelpiped$S$, the volume of the parallelpiped formed by $T(S)$ is $|\text{Det } A| \cdot$ (volume of S).

Section 4.1: Vector Spaces and Subspaces

A vector space is a non-empty set V of vectors, for which all of the following operations are applicable, for any set of vectors u, v, w and all scalars c, d:

- $u + v$ is in V
- Commutative sum: $u + v = v + u$
- Associative sum: $(u + v) + w = u + (v + w)$
- $u + 0 = u$
- $1 \cdot u = u$
- For every u in V, there is a unique vector $-u$ in $V \rightarrow u + (-u) = 0$
- cu is in V
- Scalar distribution: $c(u + v) = cu + cv$
- Vector distribution: $(c + d)u = cu + du$
- Associativity: $c(du) = (cd)u$

A subset H of a vector space V is a portion that has three properties:

- The zero vector is in H
- For any two vectors u, v in H, $u + v$ is in H
- For any vector u in H, and for any scalar c, the vector cu is in H

Example: For R^3, any plane that passes through the origin is a subspace - a subspace consisting of only the zero vector is called a zero subspace, written as $\{0\}$

For any set of vectors, the span of the linear combinations of the vectors constitutes a subspace of the space

Section 4.2: Null Spaces, Column Spaces, and Linear Transformations

For a set of homogeneous equations, where $Ax = 0$ has at least one solution, the solution set x is called the null space of the system - that is: Nul $A = \{x : x$ is in R^n and $Ax = 0\}$ - the null space is a subspace of R^n

To determine whether a vector x is in the null space of a matrix A, just check whether $Ax = 0$

Example: For a matrix $A = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix}$, $u = \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix}$ is in the null space:

$$\begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The above description involves an implicit definition: if x satisfies $Ax = 0$, then x is in the null space - defining the actual null space involves finding the solution set of $Ax = 0$

Example: For the matrix $A = \begin{bmatrix} -3 & 6 & -1 & 1 & 7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$, the null space is defined by:

$$\begin{bmatrix} -3 & 6 & -1 & 1 & 7 & 0 \\ 1 & -2 & 2 & 3 & -1 & 0 \\ 2 & -4 & 5 & 8 & -4 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

This leads to the solution set:

$x_1 - 2x_2 - x_4 + 3x_5 = 0$
$x_3 + 2x_4 - 2x_5 = 0$
Rewriting this in terms of each variable:
$x_1 = 2x_2 + x_4 - 3x_5$
$x_2 = x_2$
$x_3 = -2x_4 + 2x_5$
$x_4 = x_4$
$x_5 = x_5$

Thus, the null space is defined by x = $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}$

The vectors in this equation can be rewritten as u, v, w, leading to the solution that the null space of $A = x_2u + x_4v + x_5w$ - this vector set contains a vector for each free variable - also, the vector set u, v, w is necessarily linearly independent, because each coefficient is a free variable

The column space of a matrix A is the set of all linear combinations of the columns of A - that is: Col $A = \{b : b = Ax$ has a solution for all x in R^n - the column space is a subspace of R^m

Example: For the matrix $A = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3* & 7 & -8 & 6 \end{bmatrix}$, Col A is a subspace of R^3, and Nul A is a subspace of R^4.

To determine whether a vector x is in the column space of a matrix A, reduce $[Ax]$ to echelon form and determine whether it is consistent

Linear transformation: A function T is a linear transformation from vector space V to vector space W if:

- Every vector x in V maps to a unique vector $T(x)$ in W
- T(u+v) = T(u) + T(v)
- T(cu) = c T(u)

For a linear transformation T, the null space of the standard matrix is called the kernel of T, and the range of T is the set of all vectors in W mapped by at least one $T(x)$

Section 4.3: Linearly Independent Sets; Bases

Revisiting the definition of linear independence, it is notable that a vectors set is linearly dependent if and only if one vector is a linear combination of the others - e.g., if $p_1(t) = 1, p_2(t) = t, p_3(t) = 4 - t$, then this system is linearly dependent because $p_3 = 4p_1 - p_2$

If H is a subspace of vector space V, then an indexed set of vectors B is called a “basis” of H if B is linearly independent, and if the subspace spanned with B coincides with H - “standard basis” for R^n: a square identity matrix of size n

A set of vectors may span H but may not be a basis because one vector is the sum of the others, i.e., the vector set is not linearly independent - from this vector set, the redundant vectors may be removed to create a basis for H - for every subspace H and every vector set S that spans H, a subset of S comprises a basis for H

For a given vector set B, a basis for Nul B can be formed using the algorithm provided in section 4.2 to find the solution algorithms - since every vector in that set is associated with a free variable, that solution set is necessarily independent

For a given vector set B, a basis for Col B can be found by removing all non-pivot columns - example:

$$B = \begin{bmatrix} 1 & 4 & 0 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Discarding all non-pivot columns, a basis for Col $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

Spanning Set Theorem: For any set of vectors S that span a vector space H, if any vector v in S is a linear combination of other vectors in S, then removing v from S will lead to a reduced set that also spans H - if no vector can be removed so that S continues to span H, then S is a basis for H

Basis for Nul A: The algorithm provided in section 4.2 for finding the explicit definition of Nul A produces a linearly independent vector set; by definition, this vector set is a basis for Nul A.

Basis for Col A: This can be identified by discarding every vector of Col A that is a linear combination of the others - if A is provided in echelon form, simply remove all non-pivot columns - if not, then reduce A to echelon form, identify the non-pivot columns, and then remove those columns from the original set A

Two observations about the basis: A basis is the smallest possible spanning set for a vector space, since no columns can be removed that still span the vector space - a basis is also the largest possible spanning set for the vector space, since no vectors can be added without either adding a linear combination, or adding a new independent vector that expands the span of the column set

Section 4.4: Coordinate Systems

Unique representation theorem: Given a basis $B = b_1, b_2, \ldots$ of a vector space V, for every linear combination x in V, there exists a unique set of scalars such that x = $\mathrm{c}_1 b_1 + c_2 b_2 + \ldots$ - that is, every point in the vector space is the result of a unique linear combination of the vectors of B.

B-coordinates of x: The coordinates of x relative to the basis B are the weights $c_1, c2, \ldots$ such that $x = c_1 b_1 + c_2 b_2 + \ldots$, and are written as $\begin{bmatrix} x \end{bmatrix}_B = \begin{bmatrix} c_1 \\ c_2 \\ \ldots \end{bmatrix}$ - that is:

$B = \{b_1, b2\} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

If $\begin{bmatrix} x \end{bmatrix}_B = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$, then $x = (-2)b_1 + 3b_2 = (-2)\begin{bmatrix} 1 \\ 0 \end{bmatrix} + (3)\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$.

If B is the standard basis for the vector space, then, simply, $\begin{bmatrix} x \end{bmatrix}_B = x$.

Graphically, coordinate sets can be envisioned as graph paper, where the grid is drawn according to the vectors of the basis: draw parallel lines along each axis according to each vector

Translating coordinates into R^n: If $B = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $x = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$, $\begin{bmatrix} x \end{bmatrix}_B = (c_1, c_2) \rightarrow c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$ - the solution is $\begin{bmatrix} x \end{bmatrix}_B = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$

More generally, any basis B can be used to transform $\begin{bmatrix} x \end{bmatrix}_B$ to x, that is, $x = P_B \begin{bmatrix} x \end{bmatrix}_B$ - in this context, B is called the change-of-coordinates matrix from B to the standard basis in R^n - the inverse of the change-of-coordinates matrix can be used to convert x to $\begin{bmatrix} x \end{bmatrix}_B$, i.e., $P_B^{-1}x = \begin{bmatrix} x \end{bmatrix}_B$

Isomorphism: Any mapping between a vector space and its linear transformation according to a basis B is a one-to-one linear transformation - i.e., if any vector system V is projected onto coordinate system W by a one-to-one linear transformation, then all vector space calculations in V are accurately reflected in W and vice versa - any vector space V with a basis B containing n vectors is isomorphic to R^n

Section 4.5: Dimensions of a Vector Space

Properties of bases:

- If a vector space V has a basis B containing n vectors, then any vector set in V containing more than n vectors must be linearly dependent
- If a vector space V has a basis B_1 containing n vectors, then every basis B_n in V must contain n vectors

- If V is spanned by a finite set of vectors, then V is finite-dimensional, where the dimension of V, or dim V, is the number of vectors in any basis for V - if V is not spanned by a finite set, then V is infinite-dimensional (e.g., the space of all polynomials is infinite-dimensional)

Examples of dimensional subspaces in R^3:

- 0-dimensional subspace: Only the zero subspace
- 1-dimensional subspace: Any subspace spanned by only one vector, i.e., a line through the origin
- 2-dimensional subspace: Any subspace spanned by two linearly independent vectors, i.e., a plane through the origin
- 3-dimensional subspace: Only R^3 itself

Any subspace H of a finite-dimensional vector space V has a dimension no greater than V

Basis Theorem: If V is a p-dimensional vector space, then any linearly independent set of p elements is automatically a basis for V, and any set of p elements that spans V is a basis for V

Dimension of Col A is the number of vectors in A after removing all dependent columns, i.e., the number of pivot columns in A

Dimension of Nul A is the number of vectors in A that are removed to reduce Col A to a basis

Section 4.6: Rank

Row space: Just as the column space of a matrix A is the set of all linear combinations of the columns of A, the row space of A is the set of all linear combinations of the rows of A - row vectors work just like column vectors, but are written linearly: $r_1 = (-2, -5, 8, 0, -17)$, etc.

If two matrices A, B are row-equivalent, then their row spaces are the same - also, if B is in echelon form, then the nonzero rows of B form a basis for the row space of B, as well as the row space of A

Example: Given the matrix $A = \begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix}$, find the row space, the column space, and the null space.

Answer: $A \sim B = \begin{bmatrix} 1 & 3 & -5 & 1 & 5 \\ 0 & 1 & -2 & 2 & -7 \\ 0 & 0 & 0 & -4 & 20 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

The basis for Row A is simply the nonzero rows of A, or { (1, 3, -5, 1, 5), (0, 1, -2, 2, -7), (0, 0, 0, -4, 20) }

The basis for Col A is the reduction of A by the non-pivot rows: $\begin{bmatrix} -2 \\ 1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} -5 \\ 3 \\ 11 \\ 7 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 7 \\ 5 \end{bmatrix}$

- note that reducing A to echelon form enables an identification of which columns are linearly independent, but that the actual basis for A uses the selected columns of A, *not* from the echelon form of A

The basis for Nul A is the solution of the equation $Ax = 0$ - thus, applying the algorithm from 4.2: reduce B to reduced echelon form; rewrite as a set of linear equations; and choose the vectors formed by the free variables:

$$A \sim B \sim C = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -2 & 0 & 3 \\ 0 & 0 & 0 & 1 & -5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$x_1 + x_3 + x_5 = 0$

$x_2 - 2x_3 + 3x_5 = 0$

$x_4 - 5x_5 = 0$

$$x = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} x_3 + \begin{bmatrix} -1 \\ -3 \\ 0 \\ 5 \\ 1 \end{bmatrix} x_5$$

$$\therefore \text{Basis for Nul } A = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -3 \\ 0 \\ 5 \\ 1 \end{bmatrix}$$

Rank: The rank of a matrix A is the dimension of the column space of A, i.e., the number of pivot columns in A, or the number of non-free variables - the rank of A is always equal to the rank of A^T, i.e, the dimension of the row space of A

For any $m \times n$ matrix, the rank of A + the dimension of the null space of $A = n$ - this follows from the fact that for a matrix of n column vectors comprises pivot columns and non-pivot columns; the pivot columns define the rank, while the non-pivot columns define the dimension of the null space (see last comment in 4.5)

For an $n \times n$ invertible matrix A, all of the columns are linearly independent, and thus pivot columns; therefore, dim col $A = n$; rank $A = n$; Nul $A = \{0\}$; and dim Nul $A = 0$

Section 4.7: Change of Basis

A vector space V may be represented as a first basis B, but also a second basis C - additionally, one basis may be defined in terms of the other

Example:
$B = \{b_1, b_2\}\}; C = \{c_1, c_2\}$
$b_1 = 4c_1 + c_2; b_2 = -6c_1 + c_2$
If $\begin{bmatrix} x \end{bmatrix}_B = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, find $\begin{bmatrix} x \end{bmatrix}_C$.
Answer: We can write $\begin{bmatrix} b_1 \end{bmatrix}_C = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} b_2 \end{bmatrix}_C = \begin{bmatrix} -6 \\ 1 \end{bmatrix}$
Then, just expand: $\begin{bmatrix} x \end{bmatrix}_C = \begin{bmatrix} 3b_1 + b_2 \end{bmatrix}_C = 3 \begin{bmatrix} b_1 \end{bmatrix}_C + \begin{bmatrix} b_2 \end{bmatrix}_C$

$$\begin{bmatrix}x\end{bmatrix}_C = \begin{bmatrix}4 & -6\\1 & 1\end{bmatrix}\begin{bmatrix}3\\1\end{bmatrix} = \begin{bmatrix}6\\4\end{bmatrix}$$

The matrix $\begin{bmatrix}4 & -6\\1 & 1\end{bmatrix}$ in this example is used to transform $\begin{bmatrix}x\end{bmatrix}_B$ coordinates to $\begin{bmatrix}x\end{bmatrix}_C$ coordinates, and is identified as a change-of-coordinates matrix from B to C, or $P_{C\leftarrow B}$, such that $\begin{bmatrix}x\end{bmatrix}_C = P_{C\leftarrow B}\begin{bmatrix}x\end{bmatrix}_B$, where the columns of $P_{C\leftarrow B}$ are the equations for each component of B written in terms of C

For any two bases B, C of a vector space V, there exists a unique $n \times n$ matrix $P_{C\leftarrow B}$ such that $\begin{bmatrix}x\end{bmatrix}_C = P_{C\leftarrow B}\begin{bmatrix}x\end{bmatrix}_B$, where the columns of $P_{C\leftarrow B}$ are the columns of each component of C in B.

The inverse of $P_{C\leftarrow B}$ is $P_{B\leftarrow C}$, i.e., $P^{-1}_{C\leftarrow B}\begin{bmatrix}x\end{bmatrix}_C = \begin{bmatrix}x\end{bmatrix}_B$

Given two change-of-coordinates matrices B, C, each defined in R^n, we can find $P_{C\leftarrow B}$ by placing them side-by-side and row-reducing the left one to echelon form

Example:

$b_1 = \begin{bmatrix}-9\\1\end{bmatrix}; b_2 = \begin{bmatrix}-5\\-1\end{bmatrix}; c_1 = \begin{bmatrix}1\\-4\end{bmatrix}; c_2 = \begin{bmatrix}3\\-5\end{bmatrix}$

To find $P_{C\leftarrow B}$:

$$\left[\begin{array}{cc|cc}1 & 3 & -9 & -5\\-4 & -5 & 1 & -1\end{array}\right] \sim \left[\begin{array}{cc|cc}1 & 3 & -9 & -5\\0 & 7 & -35 & -21\end{array}\right] \sim \left[\begin{array}{cc|cc}1 & 3 & -9 & -5\\0 & 1 & -5 & -3\end{array}\right] \sim \left[\begin{array}{cc|cc}1 & 0 & 6 & 4\\0 & 1 & -5 & -3\end{array}\right]$$

$\therefore P_{C\leftarrow B} = \begin{bmatrix}6 & 4\\-5 & -3\end{bmatrix}$

Section 5.1: Eigenvalues and Eigenvectors

Eigenvectors: For a given matrix A, the transformation Ax typically yields a vector heading in a different direction - however, in some cases, a vector y exists such that Ay is in the same direction, and simply scaled with a particular value (λ), that is, $Ax = \lambda x$ - in this case, y is an eigenvector for A, and λ is an eigenvalue of A - a number is an eigenvalue of a matrix A if some vector x is an eigenvector corresponding to the eigenvalue λ for A

To determine if x is an eigenvector of A, just compute Ax and see if the resulting vector is a multiple of x

To determine if y is an eigenvalue of A, determine whether some vector x exists for which $Ax = yx$ - this can be determined as $(A - yI)x = 0$; that is, subtract the identity matrix multiple of y from A

Example: Find all eigenvectors corresponding to the eigenvalue of 7 for the matrix $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$.

Solution:

$A - 7I = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} -6 & 6 \\ 6 & -6 \end{bmatrix}$

Now find a solution for $(A - 7I)x = 0$:

$\begin{bmatrix} -6 & 6 & 0 \\ 5 & -5 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

Solutions: $x_1 = x_2; x_2 = x_2 \rightarrow x = \begin{bmatrix} 1 \\ 1 \end{bmatrix} x_2$

Therefore, all eigenvectors corresponding to eigenvalue 7 for A have the form $x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

For a particular matrix A and a particular eigenvalue λ, the set of all eigenvectors is a subspace of R^n, called the eigenspace of A corresponding to λ

For a triangular matrix, the eigenvalues are the entries on the main diagonal

What if a matrix A has an eigenvalue of zero? - then at least one nontrivial solution exists for $Ax = 0x$, which is equivalent to $Ax = 0$ - thus, any such matrix is by definition not invertible

If an $n \times n$ matrix has eigenvectors $v_1, \ldots v_r$ corresponding to distinct eigenvalues $\lambda_1, \ldots \lambda_r$, then the set $\{v_1, \ldots v_r\}$ is linearly independent

Section 5.2: The Characteristic Equation

For a particular matrix, how can the entire set of eigenvalues be determined? - eigenvalues exist only when $(A - \lambda I)x = 0$ has a non-trivial solution - this occurs only when the determinant is zero, i.e., when the matrix is non-invertible - for a 2x2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, this can be expressed as Det $\begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix} = 0$, or $(a - \lambda)(d - \lambda) - bc = 0$ - solution: expand, collect terms into a polynomial, factor, and find values for λ

Example: For $\begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}$, the eigenvalues satisfy $(2 - \lambda)(-6 - \lambda) - 9 = 0$
$\lambda^2 + 4\lambda - 21 = 0$
$(\lambda - 3)(\lambda + 7) = 0$
$\lambda = \{3, -7\}$

Review of determinants: For an $n \times n$ matrix A, if U is the echelon form of A achieved through any number of row replacements and r row interchanges (but no scaling), then the determinant of A is $(-1)^r$ times the product of the main diagonal of U - if the main diagonal includes any zeros, then the determinant is zero and the matrix is not invertible

Extension of invertible matrix theorem: An $n \times n$ matrix is invertible if and only if the determinant is not zero, and zero is not an eigenvalue of A - conversely, A only has zero as an eigenvalue if it is a non-invertible matrix

The characteristic equation: The equation Det $(A - \lambda I) = 0$ is called the characteristic equation of A - expanding and factoring this equation into a polynomial produces a characteristic polynomial of A, e.g., $(\lambda - 5)^2(\lambda - 3)(\lambda - 1) = 0$

Multiplicity: When expressed as a factored characteristic polynomial, the multiplicity of each eigenvalue is its number of occurrences - e.g., in $(\lambda - 5)^2(\lambda - 3)(\lambda - 1) = 0$, the eigenvalue 5 has a multiplicity of 2, since it occurs twice in the characteristic polynomial - this can also apply to the eigenvalue 0: $\lambda^6 - 4\lambda^5 - 12\lambda^4 = \lambda^4(\lambda - 6)(\lambda + 2) \rightarrow$ the eigenvalues are 0 (multiplicity 4), 6, and -2 (both multiplicity 1)

Similarity: Two $n \times n$ matrices A, B are similar if they have the same characteristic polynomial, i.e., the same set of eigenvalues with the same multiplicities (note: just having the same eigenvalues is not enough)

A pair of similar matrices A, B also have the following property: If A and B are similar, there exists an invertible matrix P such that $A = PAP^{-1}$ - if true, then $B = P^{-1}AP$ must also be true - this operation is called a similarity transformation

Section 5.3: Diagonalization

Given the description of similarity above, a matrix A can be expressed as PDP^{-1} where D is a diagonal matrix (i.e., only has nonzero entries along the major diagonal)

Diagonalization allows for some shortcuts for matrix operations such as computing a power of a matrix - example: given a matrix A, a diagonal matrix D, and a similarity transformation P such that $A = PDP^{-1}$, we can find a formula for any matrix A^k easily:
$A^2 = PDP^{-1}PDP^{-1} = PD(1)DP^{-1} = PD^2P^{-1}$
$A^3 = PDP^{-1}PDP^{-1}PDP^{-1} = PD(1)D(1)DP^{-1} = PD^3P^{-1}$ (etc.)
$A^k = PD^kP^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$
Multiplying this through:
$A^k = \begin{bmatrix} 2 \cdot 5^k - 3^k & 5^k - 3^k \\ 2 \cdot 3^k - 2 \cdot 5^k & 2 \cdot 3^k - 5^k \end{bmatrix}$

An $n \times n$ matrix is diagonalizable only if A has n linearly independent eigenvectors, and if the columns of P are the linearly independent eigenvectors of A - in this case, the diagonal entries of D are the eigenvalues of A corresponding to the eigenvectors of P - in this case, P is called the eigenvector basis of R^n

Four-step process of diagonalizing an $n \times n$ matrix A:

1. Find the eigenvalues of A
2. Find the eigenvectors of A, and verify that they are linearly independent - if not, the matrix is not diagonalizable

3. Construct P from the eigenvectors

4. Construct D as a diagonal matrix of the corresponding eigenvectors (in the same order as the eigenvectors are laid out in P)

(Verifying step, without having to find P^{-1}: Confirm that $AP = PD$.)

This process does not require that the eigenvalues of A are distinct: an eigenvalue of A may exist with a multiplicity greater than 1, and may have more than one eigenvector - just find the set of eigenvectors for each eigenvalue and select a linearly independent set for P, and the corresponding eigenvalues for D.

Section 6.1: Inner Product, Length, and Orthogonality

Given a pair of $n \times 1$ vectors u, v, the inner product, or dot product, is the scalar $u^T v$ - properties of inner products:

- Commutative: $u \cdot v = v \cdot u$
- Distributive: $(u + v) \cdot w = u \cdot w + v \cdot w$
- $(cu) \cdot v = u \cdot (cv) = c(u \cdot v)$
- $u \cdot u \geq 0$
- $u \cdot u = 0$ if and only if $u = 0$
- $(c_1 u_1 + c_2 u_2 + \ldots c_n u_n) \cdot w = c_1(u_1 \cdot w) + c_2(u_2 \cdot w) + \ldots + c_n(u_n \cdot w)$

Vector length: A vector with entries $v_1, \ldots v_2$ has a length $||v|| = \sqrt{v \cdot v} = \sqrt{v_1^2 + v_2^2 + \ldots + v_n^2}$ - for any scalar multiple cv, the length is $c\sqrt{v \cdot v}$

Unit vector: Any vector of length 1 - every nonzero vector has a unit vector u in the same direction and of length 1, which can be determined by dividing v by its length: $u = \frac{v}{||v||}$

The distance from u to v is dist $(u, v) = ||u - v|| = \sqrt{(u_1 - v_1)^2 + (v_2 - v_2)^2 + \ldots}$

Orthogonal vectors: Two vectors are orthogonal if $u \cdot v = 0$

- This observation stems from the fact that dist $(u, v) =$ dist $(u, -v)$
- Because this relationship is true for any vector and the zero vector, every vector is orthogonal with the zero vector
- Vectors u, v are orthogonal if they satisfy the Pythagorean theorem: $||u + v||^2 = ||u||^2 + ||v||^2$

Orthogonal complement: For a vector z and a subspace W, if z is orthogonal to every vector in W, then z is orthogonal to W - the orthogonal complement of W is the set of all vectors that are perpendicular to W, and is written as $W \perp$ - for every subspace W in $R^n, W \perp$ is also a subspace in R^n

For an $m \times n$ matrix A, the orthogonal complement of hte row space of A is the null space of A, and the orthogonal complement of the column space of A is the null space of A^T

For any two vectors u, v in R^2 or R^3 that meet with an angle θ, u $\cdot v = ||u|| ||v|| \cos(\theta)$

Section 6.2: Orthogonal Sets

Orthogonal set: A vector set $\{u_1, \ldots u_p\}$ in R^n where every pair of vectors is orthogonal: $u_i \cdot u_j = 0$ whenever $i \neq j$ - every orthogonal set S is necessarily linearly independent, and the vectors form a basis for the subspace spanned by S - for any subspace W, the orthogonal basis of W is a basis that is also an orthogonal set

Orthogonal bases are helpful because any vector in R^n can be expressed as a linear combination of the vectors of the basis, with weights that are easily calculated: $y = c_1 u_1 + c_2 u_2 + \ldots + c_p u_p$, and $c_j = \frac{y \cdot u_j}{u_j \cdot u_k}$

Example: For the orthogonal basis $S = \{u_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, u_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, u_3 = \begin{bmatrix} -\frac{1}{2} \\ -2 \\ \frac{7}{2} \end{bmatrix}\}$ of R^3, express $y = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$ as a linear combination of the vectors of S.

Solution:

$y \cdot u_1 = 11; y \cdot u_2 = -12; y \cdot u_3 = -33$

$u_1 \cdot u_1 = 11; u_2 \cdot u_2 = 6; u_3 \cdot u_3 = \frac{33}{2}$

$y = \frac{11}{11}u_1 + \frac{-12}{6}u_2 + \frac{-33}{\frac{33}{2}}u_3 = u_1 - 2u_2 - 2u_3$

Orthogonal projection: Given a vector u, a vector y can be decomposed into a sum of orthogonal components: $y = \hat{y} + z$, where $hatx$ is a multiple of u, and z is a vector that is orthogonal to u - $\hat{y}$ is known as the orthogonal projection of y onto u, or $\text{Proj}_W y$, and has length α; z is known as the component of y orthogonal to u

α is the scalar length of $\hat{y}$ projected onto u, and is calculated as $\frac{y \cdot u}{u \cdot u}$ - hence, $\hat{y} = \alpha u = \frac{y \cdot u}{u \cdot u} u$

Application: Vector decomposition is useful in physics to evaluate a force vector by breaking it down into components, in order to evaluate the component of a particular force along one or more axes

Orthonormal set: An orthogonal set of unit vectors - an orthonormal basis for a subspace W is an orthonormal set that is also the basis of W (i.e., a set of orthogonal, unit vectors that spans the subspace) - easiest example: the standard basis, $\{e_1, \ldots, e_n\}$, is an orthonormal basis for R^n - an orthonormal set can be derived from an orthogonal set simply by normalizing each vector

Properties of an orthonormal set U:

- $U^T U = I$ (useful to prove that a set is orthonormal)
- $||Ux|| = ||x||$
- $(Ux) \cdot (Uy) = x \cdot y$

Orthogonal matrix: A square, invertible matrix U where $U^{-1} = U^T$

Section 6.3: Orthogonal Projections

For a subspace W of R^n, any vector y in R^n can be decomposed into a vector in W and a vector that is orthogonal to W - the vector in W is $\hat{y}$, the projection of y on W, and the orthogonal vector, z, is simply $y - \hat{y}$

For a set of vectors $u_1, u_2, \ldots$ forming an orthogonal basis for W, the orthogonal decomposition of a vector y is $\hat{y} = \frac{y \cdot u_1}{y \cdot u_1} u_1 + \frac{y \cdot u_2}{y \cdot u_2} u_2 + \ldots$

Properties of orthogonal projections:

- Wy does not depend on the particular orthogonal basis for W: any orthogonal basis for W will generate the same $\hat{y}$ for any Y
- If y is in W, then $\text{Proj}_W y = y$
- Best approximation theorem: For a vector y and a subspace W, the orthogonal projection of y onto W is the closest point in W to y, i.e., any other point in W has a farther distance to y - another description is that $\text{Proj}_W y$ is the best approximation of y by the elements of W
- When $(u_1, \ldots)$ is also orthonormal, each $u_1 \cdot u_1 = 1$, so that part of the calculation can be disregarded, and $\text{Proj}_W y = (y \cdot u_1)u_1 + (y \cdot u_2)u_2 + \ldots$ - even easier, for an orthonormal basis, $\text{Proj}_W y = UU^T y$

Section 6.4: The Gram-Schmidt Process

Given two vectors (u, v) (not necessarily orthogonal) that span W, an orthogonal basis for W can be found as u and the component of v orthogonal to u, which is simply $v - \text{Proj}_v u = v - \frac{u \cdot v}{v \cdot v} v$

This can be extrapolated to higher dimensions through iteration: given a set of vectors t, u, v spanning W, first find u' = the component of u orthogonal to t

(*i.e.,* $u - \frac{t \cdot u}{t \cdot t} t$), then find v' = the component of v orthogonal the subspace spanned by t, u'

In order to produce an orthonormal basis, simply generate the orthogonal basis, and then scale each vector of the orthogonal basis by its length

QR factorization: Any $m \times n$ matrix A can be factored into the product of two matrices, $A = QR$, where Q is an $m \times n$ orthonormal basis for A, and R is an $n \times n$ upper triangular invertible matrix with positive entries on the diagonal - Q is constructed as above; R can be calculated as $R = Q^T A$

Section 6.5: Least-Squares Problems

Many scenarios involve an equation $Ax = b$, where A is inconsistent and yields no results - in these cases, it's helpful to find an x that makes Ax as close as possible to b, i.e., Ax is an approximation of b, such that $||b - Ax||$ is as close to zero as possible - this technique is an extension of $\text{Proj}_W y$ as the best approximation of y by the elements of W as above

The general constraint is that whatever x we choose must still be in the span of A - noting that $A^T Ax = A^T b$ yields a solution:

Example:

$$A = -\begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}; b = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$$

$$A^T b = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 & 11 \end{bmatrix}$$

$$\therefore \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

This can be solved either using row reduction, or applying $(A^TA)^{-1} = \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}$,

and $\hat{x} = (A^TA)^{-1}A^Tb = \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}\begin{bmatrix} 19 \\ 11 \end{bmatrix} = \frac{1}{2}$

Some matrices will not be invertible - i.e., at least one variable is linearly dependent on another, and at least one variable is free - this will lead to a range of solutions, which can be identified by solving the augmented matrix above

Two shortcuts may be available in special cases:

- If the vectors of A are orthogonal, then $\hat{x}$ arises as the weights on the individual vectors while generating the projection of $\hat{b}$ - i.e., if $\text{Proj}_b A = x_1a_1 + x_2a_2 + \ldots$, then $\hat{x} = \begin{bmatrix} x_1 \\ x_2 \\ \ldots \end{bmatrix}$
- If A is QR factorized, then an exact solution can be found as $x = R^{-1}Q^Tb$